

Prediction of sustained harmonic walking in the free-living environment using raw accelerometry data

Jacek K. Urbanek, Vadim Zipunnikov,
 Tamara Harris, William Fadel, Nancy Glynn,
 Annemarie Koster, Paolo Caserotti,
 Ciprian Crainiceanu, Jaroslaw Harezlak *

May 18, 2015

Abstract

Accelerometers are now widely used to measure activity in large observational studies because they are easy to wear, measure acceleration at a resolution that humans cannot report, and measure objectively the acceleration of the particular body part they are attached to. Using raw accelerometry data we propose and validate methods for identifying and characterizing walking in the free-living environment. Although existing methods can identify walking in laboratory settings, recognizing it in the free-living environment remains a major methodological challenge. We focus

*Jacek K. Urbanek is Postdoctoral Fellow, Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health. Vadim Zipunnikov is Assistant Professor, Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health. Tamara Harris is Senior Investigator Chief, Geriatric Epidemiology Section, Laboratory of Epidemiology, Demography, and Biometry, National Institute on Aging. William Fadel is PhD Student, Department of Biostatistics, Indiana University School of Medicine. Nancy Glynn is Assistant Professor, Center for Aging and Population Health, Department of Epidemiology, Graduate School of Public Health, University of Pittsburgh. Annemarie Koster is Associate Professor, Department of Health, Medicine and Life Sciences, Social Medicine, Maastricht University. Paolo Caserotti is Associate Professor, Department of Department of Sports Science and Clinical Biomechanics, University of Southern Denmark. Ciprian Crainiceanu is Professor, Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health (E-mail: ccrainic@jhspsh.edu). Jaroslaw Harezlak is Professor, Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health (E-mail: harezlak@iu.edu). This research was supported by NIH grant RC2AG036594, Pittsburgh Claude D. Pepper Older Americans Independence Center, Research Registry and Developmental Pilot Grant NIH P30 AG024826 and NIH P30 AG024827 and National Institute on Aging Professional Services Contract HHSN271201100605P. This project was supported, in part, by the Intramural Research Program of the National Institute on Aging. This research was supported by the NIH grant RO1 NS085211 from the National Institute of Neurological Disorders and Stroke, by the NIH grant RO1 MH095836 from the National Institute of Mental Health and by NIH grant 1R01 HL123407 from National Heart, Lung and Blood Institute.

on sustained harmonic walking (SHW), which we define as walking for at least 10 seconds with low variability of step frequency. Our method is robust to within- and between-subject variability, measurement configuration and device rotation. The resulting approach is fast and allows processing of a week of raw accelerometry data (~ 150 M measurements) in half an hour on an average computer. Methods are motivated by and applied to the free-living data obtained from the Developmental Cohort Study (DECOS) where 49 elderly subjects were monitored for one week (~ 300 GB of data).

Keywords: Accelerometer; movement recognition; activity, walking quantification

1 INTRODUCTION

We propose methods for estimating when and how people are walking based on high-frequency, high-throughput data obtained from wearable accelerometers. These methods produce a labeled time series at the sub-minute level indicating whether the person is walking or not. For walking periods the instantaneous step frequency and the vector magnitude are also estimated. Methods are not limited by the length of recording, type of device, or configuration set-up, which is crucial in large epidemiological studies, where hundreds or thousands of subjects wear accelerometers for weeks at a time. Such studies collect data that exhibit extraordinary levels of heterogeneity due to the natural within- and between-person variability, as well as to measuring devices that are prone to batch effects, rotations and random artifacts. Heterogeneity makes walking prediction in the natural environment much more difficult than predicting walking in the lab (Zhang et al., 2012; Xiao et al., 2014; Fortune et al., 2014; Ahanathapillai et al., 2014; Ermes et al., 2008; Mathie et al., 2004; Yoneyama et al., 2014). The differences between data collected in the lab under strict protocols and data collected in large observational studies is, indeed, dramatic. This is probably why despite tens of papers that provide algorithms for walking prediction in the lab, there is currently no working solution for the problem that really matters: predicting walking when we do not know when and if it occurs. This made us coin the term data collected “in-the-wild” to emphasize the highly unstructured nature of the data obtained during, unsupervised, free-living, activity of humans.

A major problem is that the definition of walking from the point of view of the accelerometer measurement is quite ambiguous. For example, the accelerometry signal has a fundamentally different structure for a 6 minute relaxed walk versus multiple bouts of a few seconds of walking in the kitchen interspersed among highly complex cooking activities. Moreover, laboratory studies typically collect labeled data that is consistent with the 6 minute walking (e.g. 400 meter walking (Chang et al., 2004) or simply, 6 minute walk task (Enright, 2003)) Thus, in practice it is often hard to validate walking periods that are not consistent with the laboratory standardized definition. For these reasons we focus here on sustained harmonic walking (SHW), defined as walking for at least 10 seconds with low variability of step frequency. To illustrate the complexity of the data collected

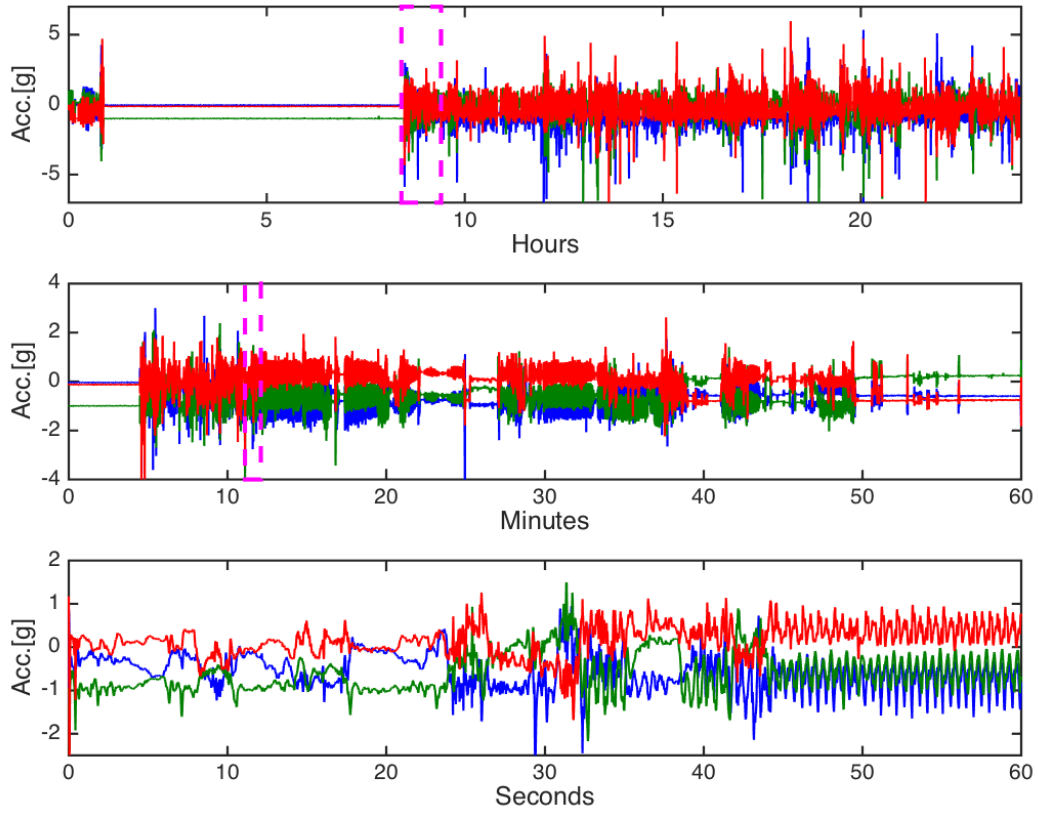


Figure 1: “In-the-wild” tri-axial acceleration signal with different colors representing data from each axis respectively. Top plot depicts data from one day of activity, middle plot from one hour and the bottom plot from one minute. Dashed-line rectangles mark zoomed-in time periods for the plots below them respectively.

“in-the-wild” Figure 1 displays three time series that represent the acceleration along three orthogonal axes of an accelerometer attached to the wrist of a person. At the most basic level, an accelerometer is an electro-mechanical device that measures acceleration along three orthogonal axes. The top panel of Figure 1 presents exactly 24 hours of the observed data, where each axis data is shown in a different color. In this example, the first observation was taken at 12 AM, while the last was taken 24 hours later, also at 12 AM. The middle plot in Figure 1 presents a particular one hour interval from 8.40 AM to 9.40 AM (indicated in the top panel as a dashed-line rectangle). The bottom panel provides an even more detailed look at the data, by displaying the one minute interval marked as a dashed-line rectangle in the middle panel (from 8.51 AM to 8.52 AM). The signal was acquired at a sampling frequency of 50Hz, i.e. there are 50 consecutive observations within each second for each of the three time series. Therefore, in practice, the number of observations per subject quickly explodes. For example, a week of accelerometry data collected at 50 Hz results in $50 * 60 * 60 * 24 * 7 = 30,240,000$ observations for each of the three axes. Thus, for even a small multi-subject observational study researchers are faced with datasets consisting of billions of observations. Our goal is to take data like the one shown in the top panel in Figure 1 and identify those periods that correspond to SHW. An example of SHW can be seen, in the lower panel of Figure 1, starting roughly around second 45. While some shorter walking periods may exist before second 45 it is impossible to validate such a claim. Thus, we will focus on methods that identify the period after second 45 as SHW. There are many scientific reasons for focusing on SHW. First, it has been reported that in-lab gait speed is related to overall survival (Studenski et al., 2011). However, little is known about how gait speed in the lab is related to enacted walking “in-the-wild” or how the amount and timing of walking periods during the day is related to survival and other health outcomes. Second, it is currently unknown what are the daily patterns of walking, how they relate to overall activity intensity and whether they are objectively associated with human health. Third, changes in timing and quality of walking during the course of the day may be associated with aging or health outcomes. Thus, we introduce methods designed for automatic identification of SHW using tri-axial body-worn accelerometers signals acquired “in-the-wild”. In addition for walking periods,

we estimate the instantaneous walking frequency (IWF) (number of steps per second) and the vector magnitude (a measure of amplitude of the acceleration signal). More precisely, starting with the original tri-axial signal $\mathbf{x}_i(t) = \{x_{i1}(t), x_{i2}(t), x_{i3}(t)\}$, where $x_{ik}(t)$ denotes the measurement for subject i at time t from the axis $k = 1, 2, 3$, we extract the vector $\mathbf{b}_i(t) = \{y_i(t), w_i(t), v_i(t)\}$, where: 1) $y_i(t)$ is a binary variable with $y_i(t) = 1$, if subject i was in SHW at time t , and 0 otherwise; 2) $w_i(t)$ is the IWF for subject i at time t ; 3) $v_i(t)$ is the corresponding vector magnitude activity count. Using the extracted vector, $\mathbf{b}_i(t)$, we can study the daily and hourly SHW percentages, the corresponding amplitude of walking accelerations, IWF variability as well as the length and timing of various bouts of SHW. These detailed summaries of walking can then be used in association studies with the demographic and clinical outcomes.

Our work was preceded by a number of other approaches see e.g. Bai et al. (2012, 2014); He et al. (2014); Troiano et al. (2014); Xiao et al. (2014). The paper by Bai et al. (2012) introduced the idea of “movelets” to build subject-specific basis functions for a variety of activities. They also provided recommendations for the conduct of large epidemiological studies collecting accelerometry data to predict activity types. Bai et al. (2014) introduced a set of metrics for human activity based on a hip-worn tri-axial accelerometer. They based these metrics on the activity time and the amplitude ratio of activity to the rest period. He et al. (2014) used movelets to analyze data from three triaxial accelerometers. Troiano et al. (2014) discuss the pressing need for a shift from count-based approaches for physical activity energy expenditure estimation to the characterization of the activities based on the features extracted from raw acceleration signals. They advocate a collaborative approach for development of analytical methods to accelerate the physical activity research. The recent technical report by Xiao et al. (2014) dealt with the inter-subject activity prediction. The authors addressed the issues of the device position variability and the normalization of the signal magnitude from different devices. While these studies have provided insights into the nature of activity type prediction, they have not been applied to highly heterogeneous data obtained from large epidemiological studies of activity “in-the-wild”.

The remainder of the paper is organized as follows. In Section 2, we introduce the accelerometry data and summarize the statistical challenges. In Section 3, we describe in

detail our method of automatic walking detection. In Section 4, we apply the proposed method to the in lab data. In Section 5, we analyze real “in-the-wild” data and we close in Section 6 with conclusions and a discussion.

2 STATISTICAL METHODS

Identification of the types of activities based on the acceleration signals is still an under-developed. We concentrate on walking, since it is one of the most common yet important activities. We focus on SHW, which is the only type of walking that we can currently consistently estimate and validate in large epidemiological studies. This may change in the future with better wearable devices or improved algorithms, but estimating SHW is important.

2.1 DATA STRUCTURE AND EDA

The motivating data were collected in the Developmental Epidemiologic Cohort Study (DECOS), a study of older individuals who are in good health. Data were collected at the University of Pittsburgh by Dr. Nancy Glynn as part of a larger study of activity in a normally aging population. The original aims of the study were: (1) to provide training sets for movement detection “in-the-wild” based on “in-the-lab” training data; and (2) to compare the activities performed (such as walking and sitting) “in-the-lab” and “in-the-wild”. We concentrate on the data from $N = 49$ individuals who have both “in-the-lab” and “in-the-wild” accelerometry measurements and associated demographic and clinical covariates. Each “in-the-lab” experiment consisted of a series of 31 activities including normal and fast walking, sedentary activities (e.g. lying still, writing and dealing cards), low-impact activities (e.g. shopping, folding towels and kneading) and higher-impact activities (e.g. vacuuming and chair stand). All these activities were performed in a laboratory in a presence of a human observer who marked the start and end of each activity.

“In-the-wild” accelerometry data activities were collected over a one week period in the natural living environment. In contrast to “in-the-lab” experiments, the activity labels for “in-the-wild” data are unavailable.

We start by exploring some portions of the 49 million-long time series for one of the DECOS subjects’ “in-the-wild” data. Four panels in Figure 2 display examples of the data that we labeled (1) resting, (2) change in position, (3) active, but not periodic (compound) and (4) locally periodic; data are shown in the left panels. The combined data shown in Figure 2 represent less than 0.04% of the data for this particular subject. The middle panels display the corresponding local means (solid lines with colors corresponding to those of the original times series) and standard deviations (shown as solid lines with a circle and colors corresponding to those of the original times series). Both the local means and standard deviations are based on a 10-second moving window. The right panels display the Fourier spectrum for each axis separately with colors corresponding to those of the raw data, respectively. The resting signal is characterized by a flat, very low amplitude time series centered around the local mean across all 3 axes (top left panel). Local averages of the resting signals are roughly constant, while the local standard deviation is also approximately constant but very close to zero relative to the size of the signal across the 60-second time interval (middle top panel). The Fourier spectrum displayed in the right-top panel of Figure 2 does not indicate any sizable periodic component, which confirms the observed nature of the data. The signal for change in position, displayed in the second row/left panel of Figure 2, is visually similar to the resting signal, with the exception of a change in the mean of two of the time series roughly between seconds 5 and 25. This change corresponds to an increase in standard deviation. These characteristics are well captured by the local mean and standard deviation of the time series displayed in the middle-second row panel of Figure 2. The Fourier spectrum has a declining peak at zero, which is characteristic of a low frequency change, associated with the observed change in the mean. Such a spectrum is characteristic of no signal variability for all but the lowest frequencies. The compound signal, displayed in the third row/middle panel of Figure 2, is characterized by large changes both in the magnitude and frequency of the signal along all 3 axes. The mean and standard deviation exhibit substantially larger time-dependent changes compared to the mean and standard deviations of the other three types of signals displayed in Figure 2. The Fourier spectrum of the compound signal (third row/right panel of Figure 2) indicates that the signal is a combination of many fundamental frequencies that seem to not be

well synchronized across the 3 axes (the peaks and valleys of the three Fourier spectra do not overlap). The locally periodic signal, displayed in the fourth row/right panel of Figure 2, is characterized by a more systematic, periodic-looking time series around quasi-constant local means. This latter characteristic of the data can be seen in the local means displayed in the fourth row/middle panel of Figure 2. Indeed, the local averages are almost indistinguishable from the local averages displayed in the corresponding panel for resting. This indicates that resting was probably “resting standing”, as the average orientation of the device is roughly the same as for walking. A different posture would result in a flat, but dramatically different local mean level. The local standard deviations are larger for the locally periodic signal than for the resting signal, as accelerations associated with the periodic movement vary a lot more along their local means. The Fourier spectrum for the locally periodic signal, displayed in the fourth row/right panel of Figure 2, shows that there is a small number of frequencies that explain the observed variability of the signal. Moreover, the spectra of signals corresponding to different axes are better synchronized than spectra for the compound signal (compare panels on the right in the third and fourth row of Figure 2). We will use this property of the local periodic signal to detect walking. Recognizing walking in complex data obtained from observational studies “in-the-wild” is difficult because: 1) the timing and length of walking periods are unknown; 2) the intensity and frequency of walking can vary substantially within-subject and dramatically between subjects; 3) compound movements can include walking and may correspond to a large number of heterogeneous behaviors; 4) orientation, location, and type of device can substantially alter the type of signal along individual axes; and 5) Fourier decompositions become noisier in shorter intervals (3 - 5 seconds).

2.2 MODELING APPROACHES

In this section, we provide the intuitive basis of the proposed accelerometry signal modeling approach. A fundamental property of SHW is that it consists of a number of steps, which are executed within a limited range of frequencies. Our approach will take advantage of this property and estimate short periods that exhibit strong periodicity around a range of possible frequencies. Frequency of movement will be allowed to continuously adapt to a

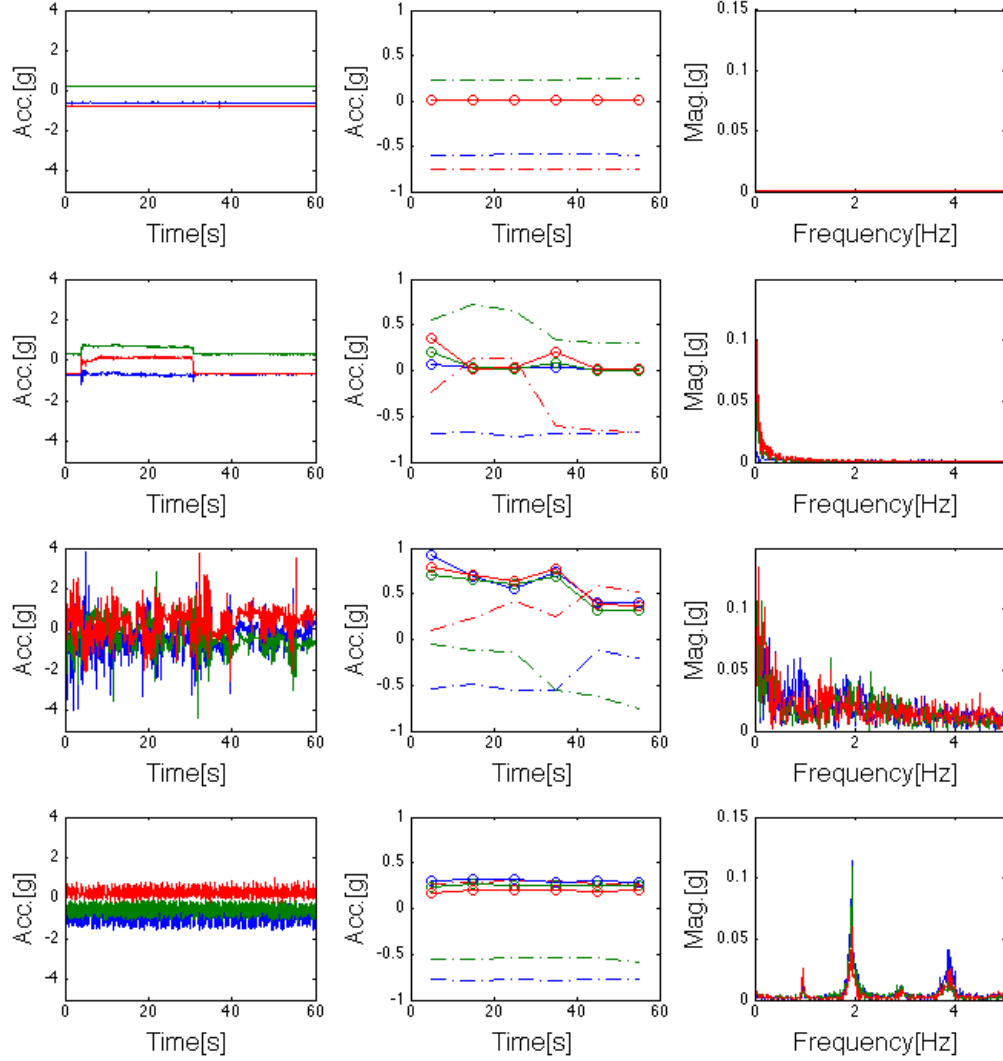


Figure 2: Panels represent four different types of signal with the units on the y-axis in earth acceleration ($g = 9.81m/s^2$) and time in seconds on the x-axis for the 3 orthogonal axes represented by different colors. Specifically, *First row, left panel* shows a portion of a resting signal, *Second row, left panel* shows a part of the change in position signal, *Third row, left panel* shows compound signal and *Fourth row, left panel* shows a locally periodic signal. Middle column displays local averages and standard deviations of the signal for the respective types of the signal and the right column displays Fourier spectra calculated for 60 second time interval.

range of walking motions both within- and between-subjects. Methods are based on Fast Fourier Transform (FFT) along each of the three axes.

The use of local FFT may be perceived as not being particularly new, though the implementation is designed and highly modified to address the size and complexity of the problem. Indeed, for experts in accelerometry data it should be obvious that a brute force approach may result in serious biases in estimating the timing and duration of walking and may not generalize well to a population of subjects. We also propose an innovative approach to using the information in the spectrum, by incorporating information at multiple of walking frequencies.

2.3 STATISTICAL CHALLENGES

A number of issues arise in modeling accelerometry data generated during SHW. First, walking is highly variable across people. Second, the device can shift or rotate relative to its original location during walking. This requires careful treatment of the signal, as such changes happen at unknown intervals and for unknown durations. The local Fourier transform is well suited for addressing those problems, as the associated changes in the accelerometry signal are typically associated with low frequency changes (see, for example, Fourier spectra for the change in position data in 2). Thus, low frequency components can be filtered out using a high-pass filter (essentially subtracting the Fourier coefficients corresponding to low frequencies).

3 METHOD FOR AUTOMATIC WALKING DETECTION

3.1 DEFINITIONS AND NOTATION

We have defined SHW as accelerometer activity signals that are performed for at least 10 seconds and are periodic with a roughly constant walking frequency. Here we propose a method for automatic recognition of SHW intervals together with estimation of the periodicity of the dominating component. The reported frequency of human walking “in-the-wild” ranges between 1.4 - 2.5 Hz (Pachi and Ji, 2005). However, we will conservatively focus on the range 1.2 - 4.0 Hz to include slow walking of older individuals and running.

We assume that any activity where the dominant frequency of movement is in this range is SHW.

Let the measured signal be $\mathbf{x}(t) = \{x_1(t), x_2(t), x_3(t)\}$, where $x_k(t)$ denotes the measurement at time index t along the axis $k = 1, 2, 3$. For notational simplicity we drop the subject index. Time is indexed from 1, which refers to the first accelerometry observation and not the wall clock. Increments are in fractions of a second. For example, time $t = 151$ corresponds to the beginning of the third second of monitoring when data are collected at 50 Hz (50 times per second).

We denote by f_0 the sampling frequency of the data expressed as number of observations per second. For example, if the acceleration signal is sampled at the $f_0=50\text{Hz}$ frequency data consist of 50 samples per second for each axis for a total of 150 observations per second for the three axes.

Next, we formally define the vector magnitude at time t as

$$vm(t) = \sqrt{x_1(t)^2 + x_2(t)^2 + x_3(t)^2}. \quad (1)$$

However, in order to compensate for the constant earth gravity we subtract 1 and define the vector magnitude count for a given period $[t - \tau/2, t + \tau/2]$ as,

$$v_\tau(t) = \frac{1}{\tau} \sum_{u=[t-\tau/2]}^{[t+\tau/2]} |vm(u) - 1|, \quad (2)$$

where τ is the window size expressed as number of sampling points. Our main goal is to estimate the vector $y(t)$, where $y(t) = 1$ corresponds to SHW and $y(t) = 0$ corresponds to non-SHW. This is achieved by employing a short-time Fourier transform (STFT) for each axis separately. The discrete Fourier transform (DFT) for axis k is $X_k(f) = \sum_{t=0}^{T-1} x_k(t) e^{-i2\pi f \frac{t}{T}}$, where f is the frequency index, t is the time index and T is the total number of observations. We define the short time Fourier transform (STFT) at time t for axis k of the acceleration signal $\mathbf{x}(t)$ as $X_k(t, f; \tau) = \sum_{u=[t-\tau/2]}^{[t+\tau/2]} x_k(u) h(u) e^{-i2\pi f \frac{u}{\tau}}$, where τ is a tuning parameter specifying the number of observations in the interval centered at t . The weights $h(u)$ are used to assign more weight to observations that are close to t . We use the Hann window which is defined as, $h(u; \tau) = 0.5[1 - \cos\{2\pi u/(\tau - 1)\}]$. The Hann window is a popular choice in applied time series analysis because it has been shown to reduce aliasing

(Harris, 1978). For this reason it has been used extensively in multiple applications including speech processing (Benesty et al., 2012), seismology (Van Herwijnen and Schweizer, 2011), vibro-acoustics (Antoni, 2005) and sleep analysis (Crainiceanu et al., 2011). The spectrum is defined as $S_k(t, f; \tau) = |X_k(t, f; \tau)|$, where $|\cdot|$ denotes the absolute value of a complex number. A first step towards the automatic walking recognition is to introduce a function that “combs” the spectrum for those frequencies that are likely to be related to walking. A comb function is defined by a fundamental frequency, say s , and the thickness of the comb “teeth”. All teeth are considered to be of the same width (same number of frequency bins) and the total number of teeth is n_m . For any frequency s , we define a neighborhood, $N_s = \{s - \frac{1}{T}, s, s + \frac{1}{T}\}$, where T is the length of the window expressed in seconds, where the local FFT is applied. Thus, N_s is the shortest frequency interval centered at s and consisting of three frequencies. Intuitively, we are using “a comb with very narrow teeth”. The comb family of functions, $C(f; s)$, indexed by s , is defined as

$$C(f; s) = \begin{cases} 1 & \text{for } f \in \cup_{l=2}^{n_m} N_{ls/2} \\ 0 & \end{cases} \quad (3)$$

While notation is a bit involved, the intuition behind the comb functions is straightforward. Consider the spectra shown in the top panel of figure 3 that correspond to a particular walking interval. The bottom panel displays the comb function corresponding to the fundamental frequency $s = 2.067$ Hz, the peak of all three spectra in the top panel (black teeth). The tooth centered at 2.067Hz also contains the frequencies 1.967Hz=2.067-1/10Hz and 2.167Hz=2.067+1/10Hz, because we are using a window of size $T=10$ s and the resulting spectral frequency is equal to $1/T = 0.1$ Hz. There are additional “comb teeth” centered at $3s/2 = 3.102$ Hz and so on. Figure 3 also displays another comb shown in red corresponding to $s = 2.81$ Hz. Note that integrating (summing) the spectrum in areas corresponding to the “black” comb will result in a higher value than for the spectrum integrated in the areas corresponding to the “white” comb, which does not match the spectral peaks. Thus, the intuition behind our proposal is relatively simple: use comb functions, multiply them with the spectrum, add these values and choose those combs that maximize the partial area under the spectrum. This is especially useful when the process is repeated millions of times. The comb function idea was inspired by the widely used comb filter of

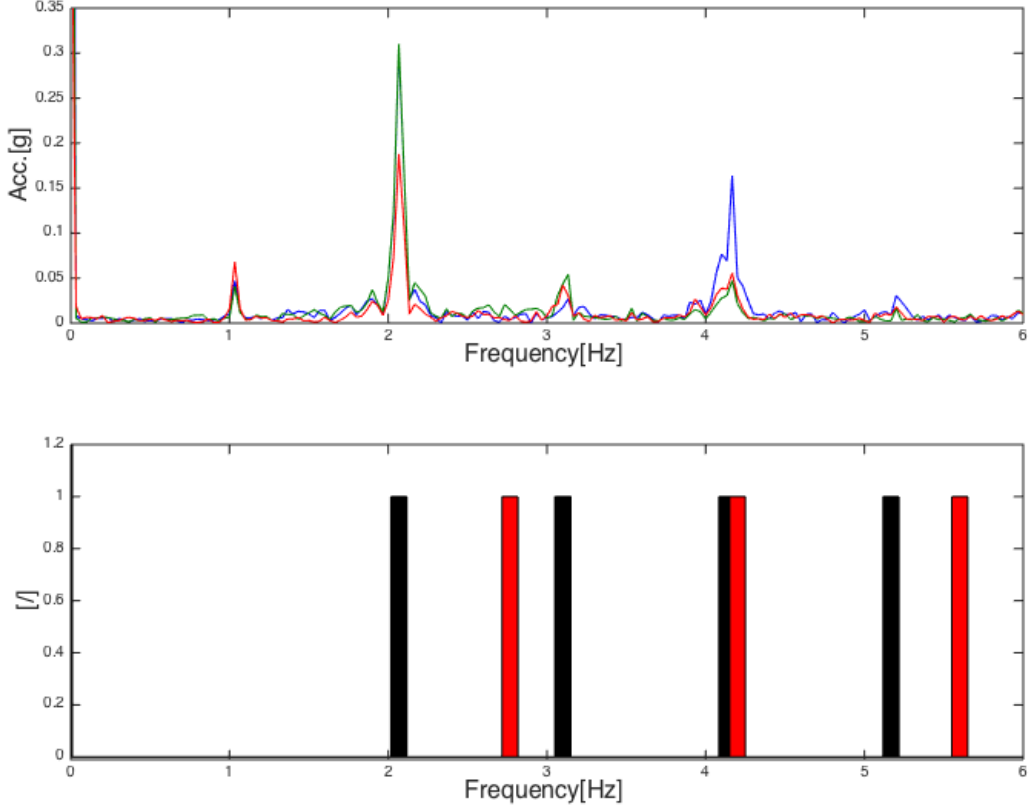


Figure 3: Top figure shows Fourier spectra of tri-axial acceleration signal of walking. Bottom figure shows “Comb” function $C(f; s)$ for $s = w(t) = 2.067$ Hz (*black*) and for $s = 2.81$ Hz (*white*). *Black* lines correspond to the spectral lines representing walking signal components while *white* lines miss those spectral lines completely.

harmonic components (Deller et al., 2000). However, in contrast to this filter that uses the entire range of spectral components, our comb uses only frequencies between s and $n_m s$. In walking recognition it is necessary to limit the number of high frequency signals, as the signal is noisy, which induces spurious high frequency spectral spikes. In simpler terms, FFT does not eliminate the noise, but re-distributes it, especially in the high frequency range.

3.2 Prediction of walking periods and its characteristics

We now provide the technical description of the SHW prediction approach. Specifically, for each axis k , we define the area under the full spectrum, $S_k(\cdot)$, as $IS_k(t) = \int_{f=0}^{f_0/2} S_k(f, t; \tau) df$ and the area under the spectrum corresponding to the comb function $C(f, s)$ as $IS_k(t, s) = \int_{f=0}^{f_s/2} S_k(f, t; \tau) C(f; s) df$. All the functions under the integral are positive and $IS_k(t) > IS_k(t, s)$ for every s and t , except for degenerate cases (e.g. when all values of the spectrum are zero). We define

$$Y_k(t, s) = \frac{IS_k(t, s)}{IS_k(t) - IS_k(t, s)}, \quad (4)$$

which is a measure of the size of the periodic content of the accelerometry signal along axis k corresponding to the comb function $C(\cdot; s)$ relative to the total variation associated with the signal along axis k .

Let $Y(t, s) = \max_k \{Y_k(t, s)\}$, which is the maximum of the fraction of the signal explained by the frequency s along the three axes $k=1,2,3$. This is the time series that is used to estimate SHW periods. For a given a threshold δ , we estimate that a person is in SHW as

$$\hat{y}_\delta(u) = \begin{cases} 1 & \text{if } \max_{s \in D_f} Y(t, s) > \delta; \\ 0, & \end{cases} \quad (5)$$

for every $u \in [t - \tau/2, t + \tau/2]$, where δ is a tuning parameter and D_f is the family of frequencies corresponding to walking. Here we use $D_f = [1.2, 4]Hz$, which is a conservative choice. Once the estimation process of $y_\delta(\cdot)$ is finished for the interval $[t - \tau/2, t + \tau/2]$, the interval is shifted by a second and the procedure is repeated. Due to the use of overlapping windows, for every 1-second intervals there will be 9 estimates of whether or not the person is walking. Based on these estimators we predict that the person is walking in that particular second if any of the 9 intervals estimated the person to be walking.

For periods that are estimated as walking the IWF index will be estimated as the frequency, s , that maximizes $Y(t, s)$. More precisely,

$$\hat{w}(t) = \begin{cases} \operatorname{argmax}_{s \in D_f} \{Y(t, s)\} & \text{for } \hat{y}_\delta(t) = 1 \\ NaN & \end{cases} \quad (6)$$

The complete description of the approach is summarized in Algorithm 1.

Algorithm 1

Input: $\mathbf{x}(t)$ - accelerometry signal, f_s - sampling frequency, T - observation time, τ - time window, δ - threshold, $s_{min} = 1.2Hz$, $s_{max} = 4.0Hz$.

Output: $y(t)$ - walking indicator, $v(t)$ - vector magnitude, $w(t)$ - IWF.

Step 1. Compute the value of the “comb” function $C(f; s)$ for each value of $s \in D_f$

Step 2. Compute the spectrum, $S_k(t, f, \tau)$, for each axis $k = 1, 2, 3$ and the area under the spectrum $IS_k(t)$ for each t

Step 3. Compute the partial area under the spectrum $IS_k(t, s)$ for each s and each k .

Step 4. Calculate the periodic content of the signal $Y_k(t, s) = \frac{IS_k(t, s)}{IS_k(t) - IS_k(t, s)}$ for each k .

Step 5. Estimate $y_\delta(t)$ with $\max_{s \in D_f} Y_k(t, s)$.

Step 6. For the times t with $\hat{y}_\delta(t) = 1$ estimate $w(t)$ by finding the s that maximizes $Y(t, s)$.

Step 7. Identify walking if any of the estimators $\hat{y}_\delta(t)$ is 1.

The properties of the proposed method depend on three tuning parameters: the short-time window length, τ , used for the Fourier transformation, the threshold δ used for estimating the cut-off point on $\max_{s \in D_f} \{Y(t, s)\}$ that indicates SHW/non-SHW and number of harmonics (comb teeth) n_m used for calculation of $IS_k(t, s)$. We evaluate the tuning parameter selection in Section 4.1 and provide recommendations for their choices based on “in-the-lab” experiment. The properties of the proposed method are studied in Section 4 where ROC curves in conjunction with the corresponding areas under the ROC curves (AUCs) are estimated for the SHW detection in “in-the-lab” data.

4 VALIDATION USING “IN-THE-LAB” DATA

To validate the method, we used data collected during the “in-the-lab” phase of the DECOS study. We use data collected on $N = 47$ elder participants, age (Q1=74, median=78, Q3=78.28), 25 males and 22 females. During the “in-the-lab” phase of the experiment all participants were asked to perform a series of physical tasks including: lying still, standing still, washing dishes, sitting still, dough kneading, dressing, folding towels, vacuuming, shopping, writing, dealing cards, standing up from a chair, walking for 20 meters, walking for 20 meters with arms crossed on the chest, fast walking for 20 meters, fast walking for 20 meters with arms crossed on the chest, treadmill walking at 1.5mph for 5 minutes, walking for 40 meters and fast walking for 400 meters. The experiment was supervised by a trained professional who recorded the times of the beginning and the end of each task. For walking prediction we focused on 400-meter walk, whereas all non-walking activities including dressing, shopping and standing up from a chair were classified as non-walking.

4.1 Selection of the tuning parameters

To apply our method to the raw accelerometry signal, several tuning parameters need to be specified: the window length, τ , the threshold δ and the number of harmonics (comb teeth) n_m . Here, we discuss their choices.

The selection of the length of the moving window τ , determines the resolution of the spectrum which is inversely proportional to the signal length; for example a value of $\tau = 10$

seconds corresponds to a frequency resolution of 0.1 Hz. There are trade-offs in choosing τ as: 1) longer windows result in more precise spectrum estimation if the time series are relatively stationary in the time window; and 2) shorter time windows are more likely to capture changes in walking frequency or changes in activity type. The optimal choice of τ depends on the unknown properties of the observed signal. Based on the empirical evaluation of “in-the-lab” data, we have found that a 10 second interval is long enough for people to maintain their IWF yet short enough to be sensitive to changes in the transition from SHW to non-SHW. Thus, in the remainder of this article, we use $\tau = 10$ seconds.

Choice of the threshold δ and the number of harmonics n_m are inter-related. The proportion of the variability ($\max_{s \in D_f} Y(t, s)$) explained is an increasing function of the number of harmonics n_m . We study δ as a function of n_m , for $n_m = 2, \dots, 17$, where the limit 17 is determined by the sampling frequency of the raw accelerometry data. We estimate the density function of $\max_{s \in D_f} Y(t, s)$ for all SHW and non-SHW periods for all subjects. The parameter δ was then estimated for each subject separately as the intersection between the subject-specific SHW and non-SHW density functions. Figure 4 displays the boxplot of the estimated δ values across subjects as a function of the number of harmonics, n_m . As expected, the median δ increases as a function of n_m , as every subject-specific δ is an increasing function of n_m . While some between-subject variability exists in the estimation of δ at every value of n_m , having a population level value simplifies the procedure considerably. For example, when $n_m = 6$ harmonics are used an estimated median value of δ at the population level is equal to 0.115. The selection of the number of harmonics, n_m , is important for the estimation of IWF, $\omega(t)$ (equation 6). In principle, we would like to utilize as many frequencies as possible without degrading IWF estimation. Indeed, a quick inspection of Figure 3, indicates that there is information about walking frequency at multiple harmonics (multiples of the fundamental frequency). If data were perfectly periodic then one would expect the harmonics to line up perfectly. However in practice, walking accelerometry data is not perfectly periodic and contains a substantial amount of noise.

To study the choice of n_m , we have calculated the IWF, $\omega(t)$, for every subject at each time point during their 400-meter walk. For every subject, we then calculated the coefficient

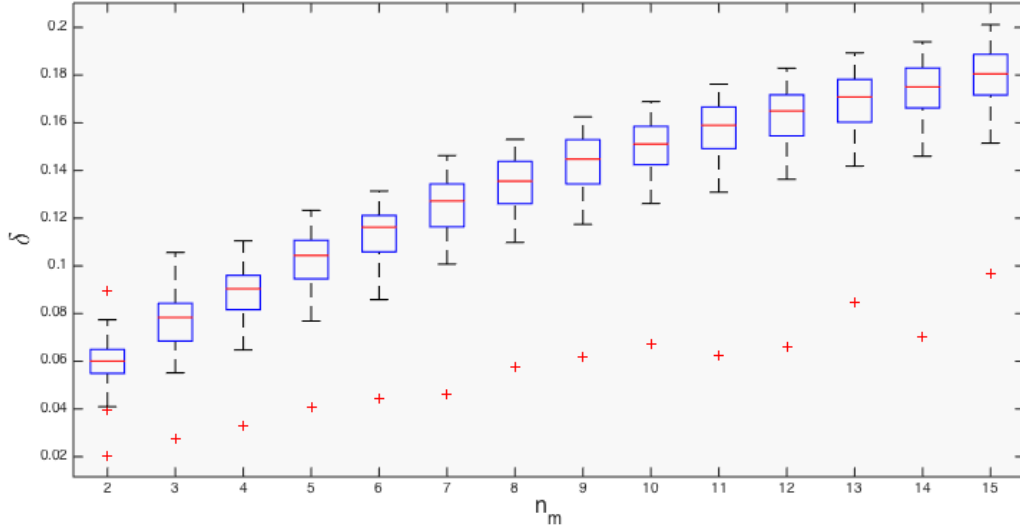


Figure 4: Boxplot representing distributions of subject-specific δ vs. n_m .

of variation of IWF during this period. Given that this is a well controlled experiment, the coefficient of variation of the IWF is expected to be small. Figure 5 displays the coefficient of variation as a function of number of harmonics for every subject (transparent lines) together with their mean (dashed line) and median (solid line). We have observed that the median coefficient of variation is relatively stable for $n_m = 2, \dots, 6$ and it starts to increase for larger values of n_m . Thus, to use as much periodic information as possible and still keep the coefficient of variation small, we selected $n_m = 6$. The road to complexity is paved with tuning parameters and we had to make some choices as well. Specifically, we used a $\tau = 10$ second with 9 second overlap between consecutive windows, $n_m = 6$ harmonics for calculating the comb function, and $\delta = 0.115$ as the threshold for percent variability explained used for predicting SHW. We also explained exactly why we made these choices and provided the technical details. We are not aware of any published paper that considered the complexity of predicting SHW “in-the-wild” and with such specific attention to details.

Many other choices could be made for the tuning parameters. For example, using a subject-specific threshold δ may seem like a more powerful approach. However, with such a choice the algorithm would need to be retrained and only work on subjects who have

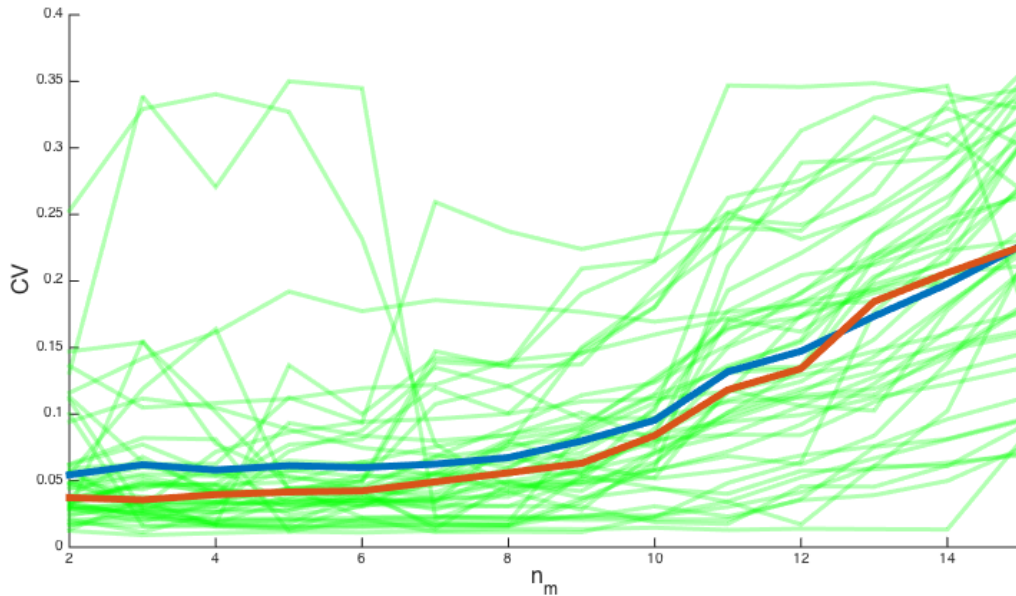


Figure 5: Coefficient of variation as a function of number of harmonics n_m . Transparent lines - subject specific, dashed line - across-subject mean, solid line - across-subject median.

carefully collected “in-the-lab” data. Moreover, in our experience, walking characteristics can change dramatically within the same person within the same day. Thus, it makes sense to identify those thresholds that do well across a population of subjects and time.

4.2 The gold standard for walking signal indices

In an ideal scenario one would need the gold standard labeling of walking and non-walking activities. Such gold standards exist for “in-the-lab” experiments, which contain the labels provided by human observers. As an example, Figure 6 displays the acceleration data with the dashed-line box indicating the portion of 400-meter-walk period identified by the human observer as a 400-meter walk. This plot shows that even gold-standard labels are actually “shifted” by about 45 seconds relative to the walking signal. Correspondingly, the period of walking activity seems to be about 45 seconds shorter than the actual signal visible as the low level of activity before the beginning of the high-amplitude, quasi-periodic signal. This can be due to multiple factors including imperfect synchronization of watches, time elapsed

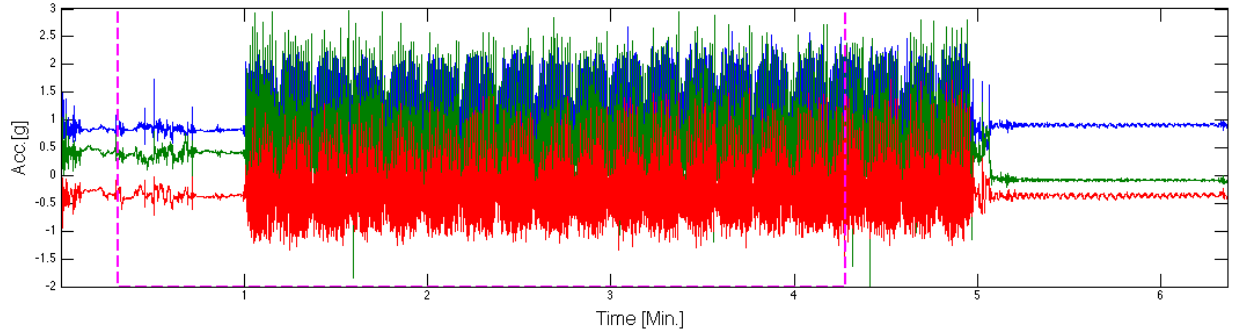


Figure 6: Tri-axial acceleration signal from 400 meters walk. Dashed-line box marks time indices given by human observer.

between the beginning and recording of the task, and basic observer error. We chose to manually inspect all “in-the-lab” data for each subject and relabel walking as periods that seem to more closely correspond to the SHW. We have found that the overlap between labels of the human observer and our improved labels was below 80% in 18 out of 49 subjects. For 29 out of the 49 subjects the overlap was above 80% (see Figure 7). Additionally, 2 out of 49 subjects did not perform the 400-meter walk task. Proper relabeling is crucial when one is interested in quantifying the performance of walking algorithms. Indeed, an algorithm designed to recognize walking patterns like the ones between minutes 1 to 5 in Figure 6, would have a hard time recognizing as walking the time series before minute 1. We contend that no algorithm should and propose to improve the human observer labels by direct inspection of the data. These improved walking labels were used as a gold standard; however, we have also calculated ROC curves and corresponding AUCs for the original labels provided by the human observer. The AUC was excellent (approximately 0.95) with a very tight distribution across subjects. In contrast, the performance relative to the copper standard was much lower. We contend that this, in fact, is a good thing, as we trust the gold standard a lot more than the copper one. Figure 8 provides the boxplots for the AUC using the human observer (copper standard) and the visual inspection plus human observer (gold standard).

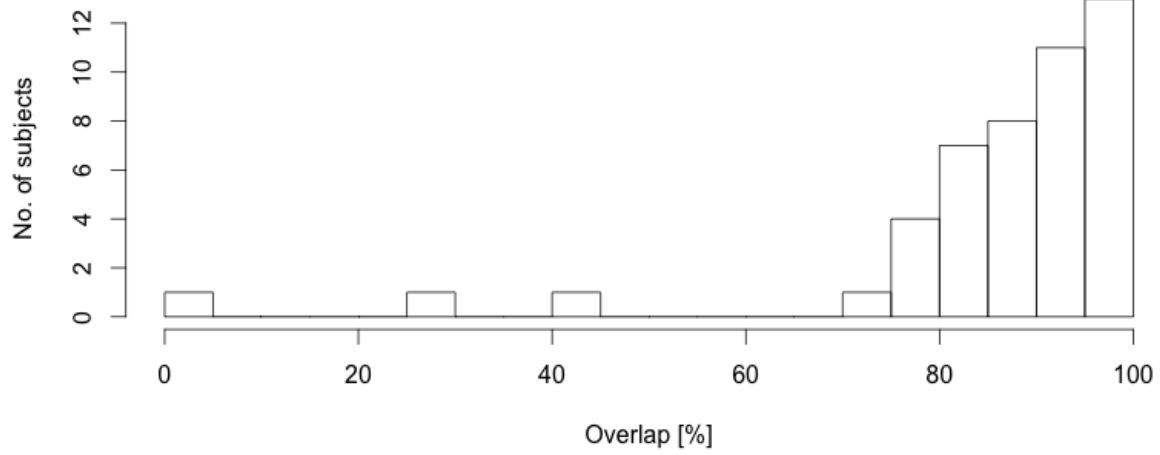


Figure 7: Summary of the overlap between times given by human observer and times corrected via visual inspection of the data.

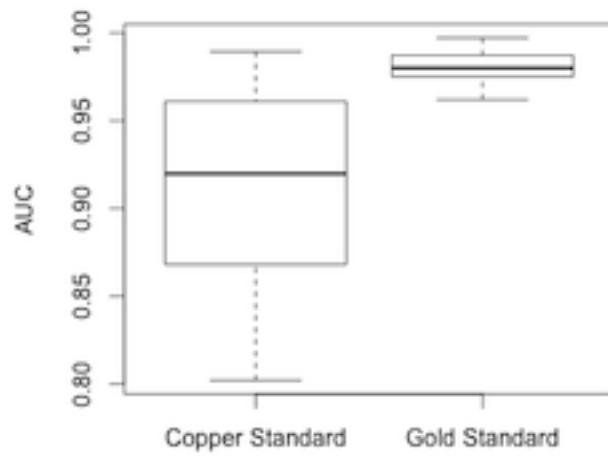


Figure 8: Boxplots for AUCs calculated for original indices (left) and for corrected indices covering whole and only walking activity (right).

5 ANALYSIS OF “IN-THE-WILD” DATA

“In-the-wild” data were collected continuously for seven days resulting in approximately 150 million observations for the three axes of each of the 49 subjects. We applied the approach described in Section 3 to estimate when SHW occurs together with the corresponding IWF and $vm(t)$. We used $\tau = 10$ and $\delta = 0.115$ and a range of possible IWF between 1.2 and 4.0 steps/second. This range is much more conservative than the range reported in the literature (Pachi and Ji, 2005) to account for the slower walking of older adults (Studenski et al., 2011) as well as for running for active individuals. SHW is estimated in time increments equal to 1 second, so consecutive windows overlap by 90% and a one second interval is declared to be SHW if any of the intervals containing it was estimated it to be SHW. The length of each walking bout is the number of consecutive time windows where walking was estimated to occur plus the window length multiplied by the overlap. For example, if walking was detected in 8 consecutive windows, the walking bout was classified as lasting for $8 + 0.9 \times 10 = 17$ seconds. It is important to state that we can not estimate a length of the walking bout that is shorter than the length of the window τ . Therefore, if walking was detected for only one window it will be classified as a bout lasting for 10 seconds.

Figure 9 displays the total SHW time (top panel) for each of the 49 participants together with the corresponding total number of walking bouts (bottom panel). Both panels present per-day averages obtained by averaging over the 7 days of activity for each subject. Participants results are sorted in decreasing order according to the estimated average total walking time. Color shading corresponds to different length of walking bouts. For example, the width of the lightest bars in the top panel of figure 9 reflects the total walking time obtained from walking bouts equal to 10 seconds. The lightest bars in the bottom panel of Figure 9 display the total number of bouts equal to 10 seconds. Results indicate that long SHW time does not necessarily indicate a large number of SHW bouts. For example, subject 1 was estimated to have the longest SHW time per day (~ 140 minutes), spread in 210 SHW bouts, which is about the third quartile of the number of SHW bouts for this population. Subject 1 had the largest number of long SHW bouts (> 60 seconds, shown in deep red). In contrast, subject 10 had an estimated average of 100 minutes of SHW

per day, that is about two thirds of the SHW time for subject 1. However, subject 10 had more than 300 SHW bouts, or about 50% SHW walking bouts than subject 1. Many of these SHW bouts correspond to short SHW walking bouts. Results also indicate that the majority of daily walking bouts for all subjects are between 10 and 30 seconds. Similarly,

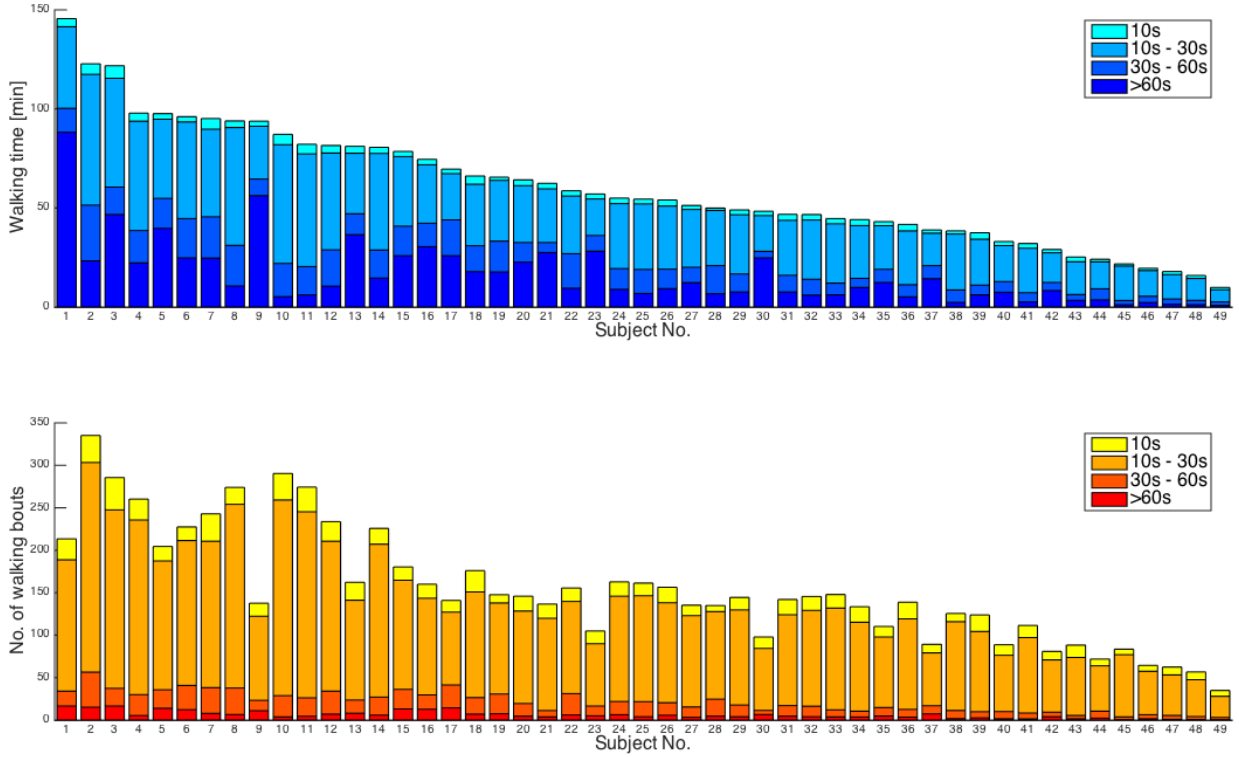


Figure 9: Bar plot presenting the total time of walking and corresponding number of bouts.

Figure 10 displays the box plots of IWF for each subjects. As in the previous plots, subjects are sorted according to total walking time, whereas the width of boxes is proportional to the number of walking bouts. The variability of IWF in this population is striking. Indeed, most subjects seem to have a wide range of IWF. For example, subject 40 had bouts of walking at 2.6 and 1.2 steps per second, which fits the currently reported range of IWF in the population. We find these plots to be deeply informative and contend that they open entire areas of research in activity that would not have been possible using currently available methodology. Figure 11 provides the lasagna plot (Swihart et al., 2010) for daily walking characteristic of all 49 subjects. The color intensity corresponds to the number of minutes of walking within a one hour window. Each row corresponds to one day and data

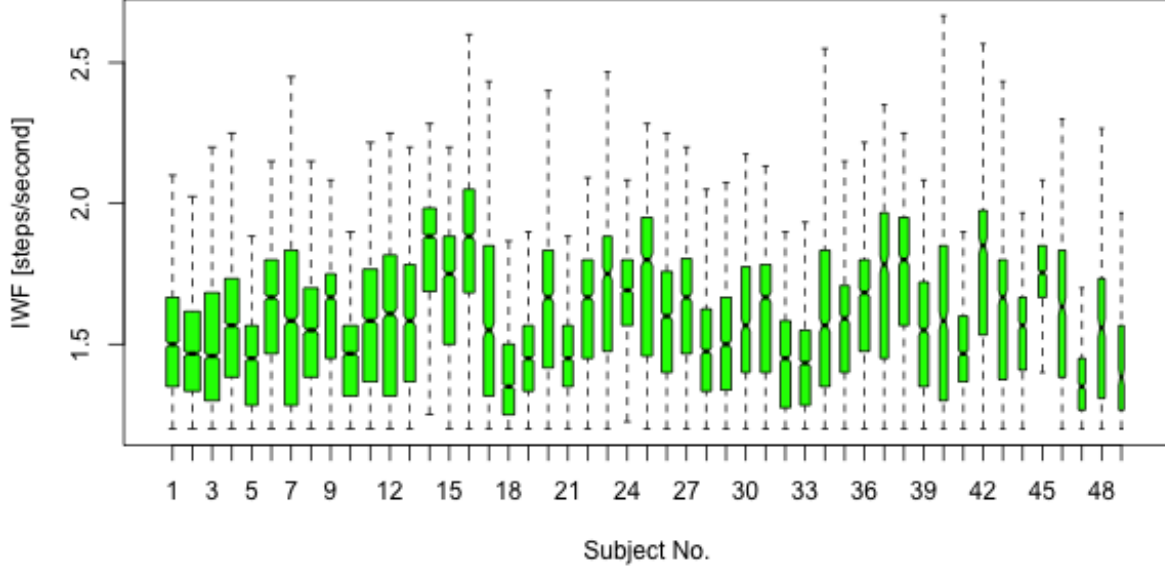


Figure 10: Box-plot presenting IWF for each participant observed “in-the-wild”. Width of boxes is proportional to the number of observations.

for individual subjects are separated by dashed lines. Subjects were sorted according to total walking time using the same ordering as in figure 9, with subjects with the highest average daily walking time shown at the top of the lasagna plot.

6 SUMMARY AND DISCUSSION

We have introduced the definition of SHW together with the prediction algorithm. We have also introduced additional quantification methods of SHW using high-density tri-axial accelerometry data. We contend that this is the first procedure designed to deal with the complexity of the data collected “in-the-wild”, the actual living environment of individuals. The method is independent of the type of device and sampling frequency of the raw data, as long as the sampling frequency is large (say $>30\text{Hz.}$). Methods have been built with the idea of being generally usable and easy to re-tune based on basic

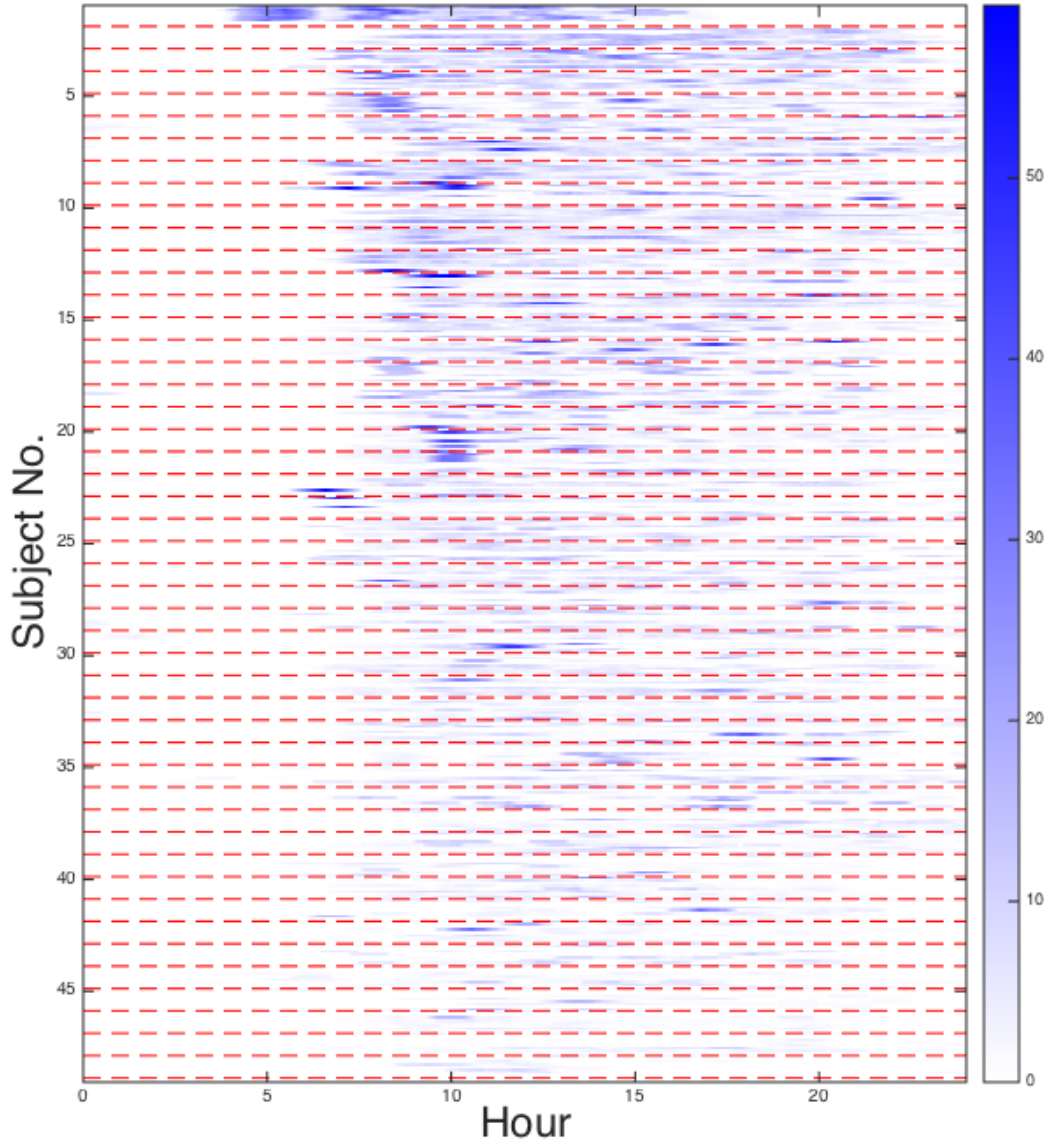


Figure 11: Lasagna plots presenting number of minutes of walking per one hour. Rows represent consecutive days. Dashed lines separate different subjects.

principles. For example, different people develop higher or lower accelerations (higher or lower amplitude signals). However, the statistic used for detecting SHW re-normalizes

data within individuals. Indeed, equation (4) indicates that decisions are based on the fraction of spectral power associated with walking and not on the total power. Thus, the normalization is intrinsic and automatic. Once SHW is identified, the vector magnitude can be calculated and the amplitude of SHW is recovered. We believe that methods work well because they are based on the quantification of the “local degree of periodicity” of all tri-axial time series as well as on sweeping the relevant frequencies using the comb function. However, it should be noted that the performance depends strongly on the assumption that SHW is a repetitive and sustained process in a particular time window. For example, it could be possible for a window of 10 seconds to contain 3-4 seconds of walking to not be recognized as walking. Thus, our method will tend to work well at recognizing sustained walking for more than 10 seconds. This problem is extremely difficult and a solution for it exceeds the scope of our current paper. Detection and analysis of a few or even single steps using extensions of the proposed method will be the subject of further studies.

References

- Ahanathapillai, V., J. Amor, M. Tadeusiak, and C. James (2014). Wrist-worn accelerometer to detect postural transitions and walking patterns. In *XIII Mediterranean Conference on Medical and Biological Engineering and Computing 2013*, pp. 1515–1518. Springer.
- Antoni, J. (2005). Blind separation of vibration components: Principles and demonstrations. *Mechanical Systems and Signal Processing* 19(6), 1166–1180.
- Bai, J., J. Goldsmith, B. Caffo, T. A. Glass, and C. M. Crainiceanu (2012). Movelets: A dictionary of movement. *Electronic Journal of Statistics* 6, 559–578.
- Bai, J., B. He, H. Shou, V. Zipunnikov, T. A. Glass, and C. M. Crainiceanu (2014). Normalization and extraction of interpretable metrics from raw accelerometry data. *Biostatistics* 15(1), 102–116.
- Benesty, J., J. Chen, and E. A. Habets (2012). *Speech enhancement in the STFT domain*. Springer.

- Chang, M., J. Cohen-Mansfield, L. Ferrucci, S. Leveille, S. Volpato, N. De Rekeneire, and J. M. Guralnik (2004). Incidence of loss of ability to walk 400 meters in a functionally limited older population. *Journal of the American Geriatrics Society* 52(12), 2094–2098.
- Crainiceanu, C. M., A.-M. Staicu, S. Ray, and N. Punjabi (2011). Statistical inference on the difference in the means of two correlated functional processes: an application to sleep eeg power spectra. *Johns Hopkins University, Dept. of Biostatistics Working Papers*, 225.
- Deller, J. R., J. G. Proakis, and J. H. Hansen (2000). *Discrete-time processing of speech signals*. IEEE New York, NY, USA:.
- Enright, P. L. (2003). The six-minute walk test. *Respiratory care* 48(8), 783–785.
- Ermes, M., J. Parkka, J. Mantyjarvi, and I. Korhonen (2008). Detection of daily activities and sports with wearable sensors in controlled and uncontrolled conditions. *Information Technology in Biomedicine, IEEE Transactions on* 12(1), 20–26.
- Fortune, E., V. Lugade, M. Morrow, and K. Kaufman (2014). Validity of using tri-axial accelerometers to measure human movement—part ii: Step counts at a wide range of gait velocities. *Medical engineering & physics*.
- Harris, F. J. (1978). On the use of windows for harmonic analysis with the discrete fourier transform. *Proceedings of the IEEE* 66(1), 51–83.
- He, B., J. Bai, V. V. Zipunnikov, A. Koster, P. Caserotti, B. Lange-Maia, N. W. Glynn, T. B. Harris, and C. M. Crainiceanu (2014). Predicting human movement with multiple accelerometers: Using movelets. *Medicine & Science in Sports & Exercise*.
- Mathie, M., B. G. Celler, N. H. Lovell, and A. Coster (2004). Classification of basic daily movements using a triaxial accelerometer. *Medical and Biological Engineering and Computing* 42(5), 679–687.
- Pachi, A. and T. Ji (2005). Frequency and velocity of people walking. *Structural Engineer* 83(3).

- Studenski, S., S. Perera, K. Patel, C. Rosano, K. Faulkner, M. Inzitari, J. Brach, J. Chandler, P. Cawthon, E. B. Connor, et al. (2011). Gait speed and survival in older adults. *Jama* 305(1), 50–58.
- Swihart, B. J., B. Caffo, B. D. James, M. Strand, B. S. Schwartz, and N. M. Punjabi (2010). Lasagna plots: a saucy alternative to spaghetti plots. *Epidemiology (Cambridge, Mass.)* 21(5), 621.
- Troiano, R. P., J. J. McClain, R. J. Brychta, and K. Y. Chen (2014). Evolution of accelerometer methods for physical activity research. *British journal of sports medicine*, bjsports–2014.
- Van Herwijnen, A. and J. Schweizer (2011). Monitoring avalanche activity using a seismic sensor. *Cold Regions Science and Technology* 69(2), 165–176.
- Xiao, L., B. He, B. Lange-Maia, N. W. Glynn, T. Harris, and C. Crainiceanu (2014). Not everybody, but some people move like you. *arXiv preprint arXiv:1404.4601*.
- Yoneyama, M., Y. Kurihara, K. Watanabe, and H. Mitoma (2014). Accelerometry-based gait analysis and its application to parkinsons disease assessment. part 1: detection of stride event. *IEEE Trans. Neural Syst. Rehabil. Eng.*
- Zhang, S., A. V. Rowlands, P. Murray, and T. L. Hurst (2012). *Physical activity classification using the GENEa wrist-worn accelerometer*. Ph. D. thesis, Lippincott Williams & Wilkins.